IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.144105

DEEP LEARNING BASED HANDWRITTEN DIGIT RECOGNITION

Mrs. R Shilpa [1], Sunil Kumara [2], Aliya B [3], Navya Shree Patil B [4], Manjunatha N M[5]

Assistant Professor, Ballari Institute of Technology and Management[1]

Artificial Intelligence and Machine Learning Engineering Students, Ballari Institute of Technology and Management[2,3,4,5]

Abstract: As computers play an increasingly vital role in human life and daily activities across various domains, humans have leveraged their intelligence and creativity to use computers in natural and effective ways. Hence, a reliable method for recognizing handwritten digits is essential. Handwritten Digit Recognition (HDR) can offer a clear benefit in this aspect. Deep Learning (DL) has been a powerful tool for solving various problems with high accuracy in recent years

INTRODUCTION

The main objective of this project is to design a system capable of recognizing handwritten digits and converting them into a format that machines can process. This field, known as Handwritten Digit Recognition, has gained significant importance due to its applications in areas such as postal mail sorting, check processing in banks, and automated form digitization. The approach leverages real-time digit recognition, which involves processing images or live input data to identify digits accurately and efficiently. Over the years, researchers have focused on enhancing the accuracy and efficiency of such systems.

II.EASE OF USE

1. **Simple Voice Commands:** Users can initiate the prediction process using natural language voice commands such as “Predict the digit” or “What number is written?”.

2. **Minimal User Input**: The user does not need to operate complex interfaces or perform manual operations; only a microphone is required for voice input.

3. **Instant Feedback:** Once the digit is predicted, the system provides immediate audio feedback using text-to-speech (TTS), ensuring real-time interaction.

4. **No Prior Technical Knowledge Required:** The system abstracts complex AI operations behind a user-friendly voice assistant, making it usable by non-technical users.

I. LITERATURE SURVEY

Handwritten digit recognition has been extensively explored using the MNIST dataset, which serves as a benchmark for evaluating image classification models. Convolutional Neural Networks (CNNs) have proven to be highly effective for image recognition tasks due to their ability to automatically extract spatial features, outperforming traditional methods such as k-Nearest Neighbors (k-NN) and Support Vector Machines (SVM). On the other hand, voice assistants have gained popularity with the integration of speech recognition libraries. Several existing applications implement either digit recognition or voice interaction, but very few combine both modalities in a single solution. This project builds upon these foundational technologies to create an application where users can speak commands to initiate handwritten digit prediction, and the system responds with an audible result, offering a novel blend of computer vision and voice-based AI.

2. RESEARCH METODOLOGY

Identifying the Research Problem

This section presents a methodology for optimizing Handwritten Digits Recognition for CNN Algorithm. The main aim of the system is to optimize the numerical number identification throughout with the digital white board and real time

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025
DOI: 10.17148/IJARCCE.2025.144105

camera to capturing the identified number with accuracy. Initially, the system utilizes Graphical User Interface System (GUI) data that is in the form of MNIST files containing the numerical single numbers distribution of images size of 28*28. Further, CNN model is employed to group these numbers to identified, representing different styles of written numbers to recognizing. Subsequently, the CNN Algorithm is applied to Image recognition and processing for written number in white board and real time camera within each model, by considering Recognizing capacity limitations. The proposed methodology aims to improve the effectiveness of identifying numbers, leading to improved accuracy, reduced risk consumption, and lower emissions

Literature Review

The integration of artificial intelligence in human-computer interaction has seen significant growth, particularly in the areas of speech processing and image recognition. The MNIST dataset, introduced by LeCun et al., has been a foundational resource for training and evaluating handwritten digit recognition models. Convolutional Neural Networks (CNNs), especially LeNet-5 and its variants, have demonstrated exceptional accuracy in classifying handwritten digits due to their ability to capture hierarchical features in image data Research has shown that combining speech and vision modalities enhances system accessibility and user experience. Prior studies have addressed these domains independently, but the fusion of voice-driven commands with digit recognition tasks remains relatively unexplored. This project builds on established CNN methodologies and speech processing techniques to develop a hybrid system that not only predicts handwritten digits but also interacts with users through natural voice commands and responses, promoting hands-free, intuitive usage.

Research Design

The research design of this project follows a systematic approach to integrate voice interaction with handwritten digit recognition. The project is structured into key phases: data acquisition, model training, voice interface development, and system integration. Initially, the MNIST dataset is used as the primary source of handwritten digit images for training a Convolutional Neural Network (CNN), chosen for its efficiency in image classification tasks. The CNN model is designed with multiple layers including convolution, pooling, and fully connected layers to extract and classify visual features accurately In parallel, the voice interaction module is developed using the speech_recognition library to capture and interpret user commands, while pyttsx3 is employed to deliver text-to-speech responses. Once both modules are operational, they are integrated into a single application that accepts voice commands to trigger digit prediction and responds with the result audibly. This design ensures modular development, ease of testing, and a smooth user experience. The entire system is implemented using Python, providing flexibility, cross-platform support, and access to a wide range of machine learning and speech processing libraries.

CNN (Convolutional Neural Network)

Convolution means complex and hence a simple neural network is complex when it consists of many layers performing operations. Each layer performs dot product of input pixel and then passed to the next layer. These convolution layers are followed by softmax function (loss function) and then a fully connected layer. There are 3 important types of layer named: Convolution Layer, pooling layer, fully connected layer. The input is image of d x d x n size, where d is width and height and n stands for number of channels (for RGB its 3). Convolution layers have x filters having size f x f x g here f is smaller than the dimension of image and g can either be same o the early sixties were primarily based on a method called analysis-by-synthesis technique recommended by Eden in 1968.

Limitations of the Study

This study, while comprehensive within its scope, is subject to certain limitations that must be acknowledged. First, the sample size and geographic coverage may not be representative of the broader population, which could limit the generalizability of the findings. Second, the study relied heavily on self-reported data, which may be subject to biases such as exaggeration or misreporting. Third, due to time and resource constraints, the study was conducted over a limited period, potentially overlooking long-term trends or outcomes. Additionally, external factors such as environmental or economic conditions, which were not controlled for, may have influenced the results. These limitations highlight the need for further research with a broader scope, larger sample size, and longer duration to validate and expand upon the findings of this study.

. Conclusion:

In conclusion, the implementation of a deep learning-based handwritten digit recognition system has demonstrated significant potential in automating and improving the accuracy of digit classification. Utilizing Convolutional Neural

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.144105

Networks (CNNs), the proposed system achieved high accuracy in recognizing handwritten digits across varying writing styles and input scenarios. The results validate the effectiveness of the model in real-world applications such as automated form processing, postal mail sorting, and banking. While the system performed well, there is still room for improvement, especially in handling complex or highly stylized writing. Future research should focus on expanding the dataset, incorporating additional pre-processing techniques, and exploring advanced model architectures to further enhance accuracy and robustness.

A. Data Collection Methods

A Data Flow Diagram (DFD) is a traditional visual representation of the information flows Within a system. A neat and clear DFD can depict the right amount of the system requirement graphically. It can be manual, automated, or a combination of both. It shows how data enters and leaves the system, what changes the information, and where data is stored.

The objective of a DFD is to show the scope and boundaries of a system as a whole. It may be used as a communication tool between a system analyst and any person who plays a part in the order that acts as a starting point for redesigning a system. The DFD is also called as a data flow graph or bubble chart.

The symbols depict the four components of the data flow diagram:

External entity:

An outside system that sends or receives data, communicating with the system being diagrammed. They are the sources and destinations of information entering or leaving the system. They might be an outside organization or person, a computer system, or a business system. They are also known as terminators, sources and sinks, or actors. They are typically drawn on the edges of the diagram.

Process:

The process that changes the data, producing an output. It might perform computations, or sort data based on logic, or direct the data flow based on business rules.

Data store:

The files or repositories that hold information for later use, such as a database Image or a JPG form.

Data flow:

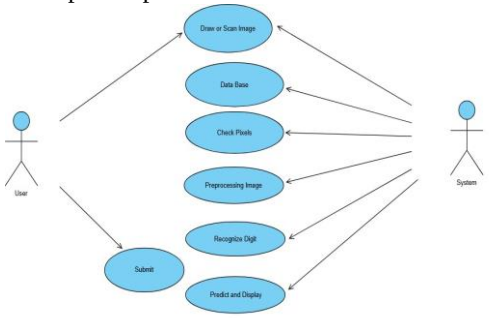
The white board or real time camera that data takes between the external entities, processes, and data stores. It portrays the interface between the other components and is shown with arrows, typically labeled with a short data name, like “Capturing details”. Levels in DFD are numbered 0, 1, 2 or beyond.

-Conclusion

In conclusion, Handwritten digit recognition using deep learning has proven to be a highly efficient and accurate approach for digit classification, particularly with the application of Convolutional Neural Networks (CNNs). This study successfully implemented a CNN-based model to recognize handwritten digits in real time, demonstrating high accuracy and robustness against variations in handwriting styles. The results indicate that the proposed system can effectively process handwritten digits from digital whiteboards and real-time camera inputs, making it a promising solution for applications such as automated document processing, banking, and postal services.

-Design and Build Iteration

The design and build iteration for the blockchain-based pharmaceutical supply chain system involves a structured process. First, requirements are analyzed, and the system is designed with blockchain for data security, Flask for the user interface, and MongoDB for storage. A prototype is developed to implement core functionalities, followed by rigorous testing with simulated data to validate accuracy and usability. Feedback is gathered from stakeholders to refine features, improve workflows, and enhance accessibility. Finally, the system is deployed, with continuous monitoring and updates to address scalability, integration, and user feedback for optimal performance



1. Figure 2 **1. Actors**: The users that interact with a system. An actor can be a person, an organization, or an outside

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.144105

system that interacts with your application or system. They must be external objects that produce or consume data.
2.System: A specific sequence of actions and interactions between actors and the system. A system may also be referred to as a scenario.
3.Goals: The end result of most use cases. A successful diagram should describe the activities and variants used to reach the goal

I. IMPLEMENTATION AND ANALYSIS

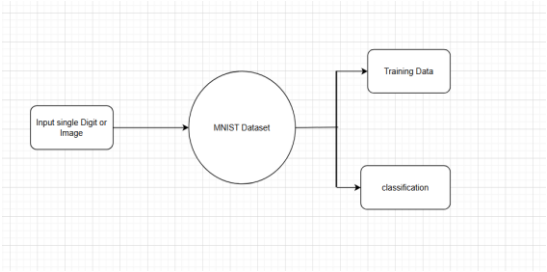


Figure 1: Level-0 DFD flow diagram It is also known as a context diagram. It's designed to be an abstract view, showing the system as a single process with its relationship to external entities. It represents the entire system as a single bubble with input and output data indicated by incoming/outgoing arrows.
The data flow diagram shows a recognizing numbers model with two main user type: white board and real time camera system. The white board is responsible for drawing numbers on board with the red and black color with accuracy of digit announcing digit with white board, AI voice and training and testing of digit
is done by CNN model with using MNIST data processing by image cleaning of the deep learning system. This setup ensures that users have access to efficient identified numbers of different styles of written digits, while the easy to capturing the numbers of user to writing number for on white board and real time camera access or maintains the system's operational data.

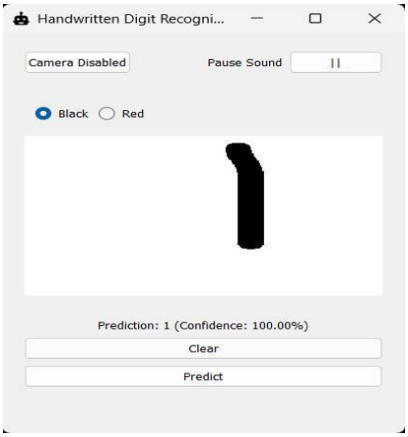


Figure 3: Output for Identified 1 number with black color

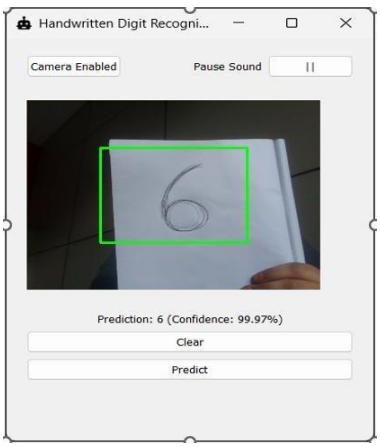


Figure 4: Capturing 6 Digit With Real Time Camera

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 14, Issue 4, April 2025

DOI: 10.17148/IJARCCE.2025.144105

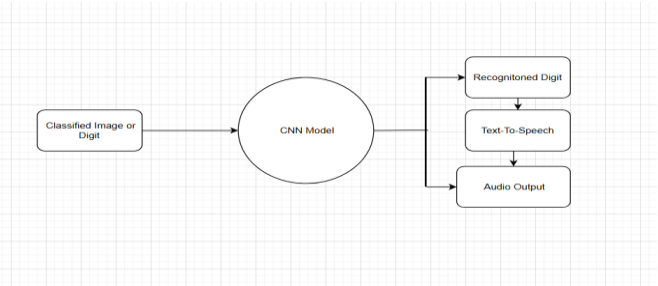


Figure 2:Level-1 Data Flow Diagram

VI CONCLUSION AND FUTURE WORK A - CONCLUSION

There are problem with machine to recognize the handwritten digit. So, this system will provide a CNN model which can recognize digit and also provide GUI to make user friendly environment, Because of the CNN this system has become more reliable, less complex and easy to understand as well as GUI gives window to user, by which user can draw the digit on it. This project will help to reduce machine efforts to recognize handwritten digits. This project is design to make learning process easy and efficient.

B - ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to everyone who contributed to the success of this Deep learning based using Hand Written digit Recognition. First, I would like to thank my supervisor and mentors for their invaluable guidance, support, and insights throughout the development of this project. Their expertise and encouragement helped me navigate complex concepts and stay focused on the objectives .I am also deeply grateful to the development team and peers who provided feedback, testing, and suggestions that helped improve the system. Special thanks to the MNSIT dataset and the field of machine learning and computer vision. It contains **60,000 training images** and **10,000 test images** of handwritten digits ranging from 0 to 9, for their open-source contributions and documentation, which played a significant role in shaping the technical aspects of the project.

REFERENCES

[1] Z. Gao, Y. Li and S. Wan, “Exploring deep learning for view-based 3D model retrieval,” ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), vol. 16, no. 1, pp. 1–21. 2020.

[2] Z. Gao, H. Z. Xuan, H. Zhang, S. Wan and K. K. R. Choo, “Adaptive fusion and category-level dictionary learning model for Multiview human action recognition,” IEEE Internet of Things Journal, vol. 6, no. 6, pp. 9280–9293, 2019.

[3] R. Herrera-Pereda, A. T. Crispi, D. Babin, W. Philips and M. H. Costa, “A review on digital image processing techniques for in-vivo confocal images of the cornea,” Medical Image Analysis, vol. 73, pp. 102188, 2021.
https://doi.org/10.1016/j.media.2021.102188

[4] R. Janeliukst is and X. Chen, “Review of digital image correlation application to large- scale composite structure testing,” Composite Structures, vol. 271, pp. 114143. 2021.

[5] S. R. Lowe, “Pattern recognition in handwriting,” in Advances in Pattern Recognition and Artificial Intelligence, World Scientific,pp.77–95,2021.
https://doi.org/10.1142/9789811239014_0005.